Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

W.H. Hundsdorfer

Unconditional convergence of some Crank-Nicolson LOD methods for initial-boundary value problems

Department of Numerical Mathematics Report NM-R9012 July

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Research (N.W.O.).

Unconditional Convergence of some Crank-Nicolson LOD Methods for Initial-Boundary Value Problems

W.H. Hundsdorfer
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

Abstract. In this paper convergence properties are discussed for some locally one-dimensional (LOD) splitting methods applied to linear parabolic initial-boundary value problems. We shall consider unconditional convergence, where both the stepsize in time and the meshwidth in space tend to zero, independently of each other.

AMS (MOS) 1980 Subject Classifications: 65M10, 65M15, 65M20.
Keywords: Numerical analysis, initial-boundary value problems, LOD methods, unconditional convergence.

1. Introduction

In this paper the accuracy of some simple splitting methods will be analysed. The methods are used for the numerical solution of initial-boundary value problems for partial differential equations (PDEs) in 2 space dimensions. Discretization in space of such PDE problems leads to a large system of ordinary differential equations (ODEs)

$$\dot{u}(t) = F(t,u(t)) \quad (0 \leq t \leq T), \quad u(0)=u_0 \tag{1.1}$$

where the vector function F contains discretized space derivatives. It is often possible to decompose F into two more simple functions F_1 and F_2,

$$F(t,v) = F_1(t,v) + F_2(t,v). \tag{1.2}$$

Standard implicit methods to approximate (1.1) require the solution of large systems of algebraic equations involving the whole function F. A well-known method is the implicit midpoint rule

$$u_{n+1} = u_n + \tau F\left(t_n + \frac{1}{2}\tau, \frac{1}{2}u_n + \frac{1}{2}u_{n+1}\right) \quad (n=0,1,2,...). \tag{1.3}$$

For linear problems, (1.3) is also often called the Crank-Nicolson method in PDE literature. The vectors u_n approximate the exact solution u of (1.1) at $t_n = n\tau$ with $\tau > 0$ the stepsize in time. Method (1.3) is of 2-nd order in the classical ODE sense.

In terms of computational effort it can be more attractive to exploit the splitting (1.2). In this paper we shall consider some locally one dimensional (LOD) methods, where the step $u_n \mapsto u_{n+1}$ is performed in two stages, in each of which only one of the functions F_1 or F_2 is used. The best known LOD method is based on the backward Euler method, see Yanenko [16]. To achieve 2-nd order, Yanenko also derived LOD methods based on the Crank-Nicolson method. In this paper we primarily consider

Report NM-R9012
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

$$u_{n+\frac{1}{2}} = u_n + \tau F_1(t_n + \frac{1}{4}\tau,\ \frac{1}{2}u_n + \frac{1}{2}u_{n+\frac{1}{2}}), \tag{1.4a}$$

$$u_{n+1} = u_{n+\frac{1}{2}} + \tau F_2(t_n + \frac{3}{4}\tau,\ \frac{1}{2}u_{n+\frac{1}{2}} + \frac{1}{2}u_{n+1}) \tag{1.4b}$$

for $n \geq 0$. The vector $u_{n+\frac{1}{2}}$ is an intermediate vector, only used for the internal computation. If F_1 and F_2 have a more simple structure than F, the computation of u_{n+1} from (1.4) can be done more efficiently than from (1.3). However, the LOD method (1.4) will have 2-nd order only if F_1, F_2 are linear and commuting; in more general situations it will be merely be of 1-st order, due to lack of symmetry, see [4]. Below we will consider two modifications which restore symmetry and 2-nd order.

Denote (1.4) as

$$u_{n+1} = \Phi_{1,2}(u_n) \quad (n=0,1,2,...) \tag{1.5}$$

where the order of the indices 1,2 refers to the fact that first F_1 was used and subsequently F_2. In this paper we shall call (1.5) the *basic scheme.*

Symmetry can be restored by interchanging after each step F_1 and F_2. This idea, which can be found in [9], leads to the following modification

$$u_{n+1} = \Phi_{1,2}(u_n), \quad u_{n+2} = \Phi_{2,1}(u_{n+1}) \quad (n=0,2,4,...) \tag{1.6}$$

which will be called here the *sequentially alternating scheme.* This modification indeed has again 2-nd order, and it requires the same amount of computational work as (1.5).

At first sight the scheme (1.6) seems superior to (1.5). This conclusion, however, will turn out not to be justified. The reason is that the classical order concept for ODEs to which we referred to until now, is the order of consistency/convergence for *nonstiff* ODEs where F satisfies a Lipschitz condition with a moderate Lipschitz constant L and τL is assumed to be sufficiently small. In our situation, where (1.1) originates from a PDE problem, the Lipschitz constant L will contain negative powers of the meshwidth in space h. As a consequence, L will be very large for fine space grids and the classical convergence theory cannot be applied. In fact, the order of the discretization errors in time may be affected by small meshwidths, a phenomenon called *order reduction.* We will see that order reduction destroys the favourable convergence properties that (1.6) has for nonstiff ODEs.

An other modification of (1.5), due to Swayne [13], reads

$$u_{n+1} = \frac{1}{2}\Phi_{1,2}(u_n) + \frac{1}{2}\Phi_{2,1}(u_n) \quad (n=0,1,2,...). \tag{1.7}$$

We shall refer to (1.7) as the *parallel alternating scheme.* Its order of consistency/convergence in the classical ODE sense is also 2. The results in [13] might give the impression that this scheme does not suffer from order reduction. It will be shown that this is not entirely correct: the order of the *local* discretization errors is reduced for small $h>0$, but it will also be shown that in the transition from local to *global* errors this reduction is annihilated due to damping and cancellation effects.

The accuracy analysis for the above LOD methods will be performed on simple linear heat-flow problems with a source term and Dirichlet boundary conditions. For space discretization standard finite differences on a rectangular grid Ω_h are considered, h being the meshwidth in space. To obtain a full convergence analysis we let τ and h tend to zero simultaneously. The effect of $h\downarrow 0$ on the orders in time are summarized by the following tabels. Only for fixed $h=h_0>0$ the classical ODE theory applies.

method	(1.5)	(1.6)	(1.7)
$h=h_0$	1	2	2
$h\downarrow 0$	¼	¼	5/4

TABLE 1.1. Orders of consistency in time

method	(1.5)	(1.6)	(1.7)
$h=h_0$	1	2	2
$h\downarrow 0$	1	[¼,½]	2

TABLE 1.2. Orders of convergence in time

The entry [¼,½] for method (1.6) in the last table means that the order of convergence is in the interval [¼,½]. These results are valid for L_2-norms. (The convergence of method (1.6) in the maximum norm is ever worse, see Table 5.3.)

The fact that LOD schemes may suffer from order reduction for $h\downarrow 0$ was mentioned already in [16] for a heat equation without source term but with boundary conditions varying in time. For such problems boundary correction techniques to restore the order of consistency are known for the basic scheme (1.5), see [8]. In some counterexamples we will see that the order reduction is also present for homogeneous boundary conditions if there is a source term in the differential equation. Since many practical problems can be modelled with homogeneous boundary conditions but with a function F which is nonlinear, or at least affine, this point should be taken into account in the derivation of boundary correction techniques. Observe from Table 1.2 that as far as the order of convergence is concerned such corrections are not necessary for (1.5), (1.7); for these two schemes the orders of convergence in time are not reduced as $h\downarrow 0$. Still boundary corrections may be useful to obtain smaller error constants, see for instance the numerical results in [12] for the LOD method based on the backward Euler method which is convergent with order 1 for both cases h fixed and $h\downarrow 0$, [10].

Although we shall deal in this paper only with simple linear problems with two space dimensions, it should be noted that the LOD schemes are stable for much more general problems, nonlinear and with arbitrary many space dimensions, see [14].

The linear model problem used for the analysis of the LOD methods is described in Section 2. The accuracy analysis of the methods can be found in the Sections 3, 4 and 5. This analysis is closely related to the one given in [6] for the Peaceman-Rachford ADI method. Method (1.5) is included in these sections, in spite of the fact that this method is known already not to have 2-nd order, not even for fixed ODEs. It will turn out that the convergence proof for (1.5) contains some basic ideas and derivations which are also applicable to the other methods.

It also possible to construct a scheme of the type (1.4) starting from the trapezoidal rule instead of the implicit midpoint rule, or to use in (1.4) not the time levels $t_n+\frac{1}{4}\tau$, $t_n+\frac{3}{4}\tau$ but only $t_n+\frac{1}{2}\tau$, for example. In Section 6 it will be shortly discussed to what extent such modification would affect the convergence results. Section 7 contains some general concluding remarks.

2. Preliminaries

2.1. The model problem

We shall consider application of the LOD methods to parabolic model problems on the unit rectangle $\Omega=(0,1)^2$ with $0\leqslant t\leqslant T$,

$$\frac{\partial}{\partial t}\mathbf{u}(x,y,t) = \Delta\mathbf{u}(x,y,t) + \mathbf{g}(x,y,t) \quad \text{on } \Omega, \tag{2.1}$$

$$\mathbf{u}(x,y,t) = \mathbf{u}_\Gamma(x,y,t) \quad \text{on } \Gamma=\partial\Omega, \quad \mathbf{u}(x,y,0) = \mathbf{u}_0(x,y) \quad \text{on } \Omega\cup\Gamma.$$

let Ω_h be the spatial grid $\{(x_i,y_j)\colon x_i=ih,\ y_j=jh,\ 1\leqslant i,j\leqslant m\}$ with $h=1/(m+1)$. Gridfunctions on Ω_h will be identified in the usual way with vectors in $\mathbb{R}^M, M=m^2$, assuming row-wise ordering on Ω_h. Thus, $w:\Omega_h\to\mathbb{R}$ will also be written as

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} \in\mathbb{R}^M \text{ with } w_j = \begin{bmatrix} w_{1j} \\ w_{2j} \\ \vdots \\ w_{mj} \end{bmatrix} \in\mathbb{R}^m, \quad w_{ij} = w(x_i,y_j).$$

Standard 2-nd order finite difference discretization of the Laplacian Δ on Ω_h leads to the semi-discrete system

$$\dot{u}(t) = Au(t) + f(t) \quad (0\leqslant t\leqslant T), \quad u(0) = u_0. \tag{2.2}$$

The vector $u(t)\in\mathbb{R}^M$ has components $u_{ij}(t)\approx\mathbf{u}(x_i,y_j,t)$, and $A=A_1+A_2$ with $A_1,A_2\in L(\mathbb{R}^M)$ approximating $\partial^2/\partial x^2$ and $\partial^2/\partial y^2$, respectively, given be the stencils

$$A\cong h^{-2}\begin{bmatrix} & 1 & \\ 1 & -4 & 1 \\ & 1 & \end{bmatrix}, \quad A_1\cong h^{-2}\begin{bmatrix} & 0 & \\ 1 & -2 & 1 \\ & 0 & \end{bmatrix}, \quad A_2\cong h^{-2}\begin{bmatrix} & 1 & \\ 0 & -2 & 0 \\ & 1 & \end{bmatrix}.$$

Further, $f(t)=b(t)+g(t)$ where $g(t)$ is the restriction of $\mathbf{g}(x,y,t)$ to Ω_h and $b(t)=b_1(t)+b_2(t)$ emanates from the boundary conditions, $b_1(t)$ having components $h^{-2}\mathbf{u}_\Gamma(x\pm h,y,t)$ on the gridpoints adjacent to the vertical boundaries and $b_2(t)$ with components $h^{-2}\mathbf{u}_\Gamma(x,y\pm h,t)$ near the horizontal boundaries. For the source term $g(t)$ splittings $g(t)=g_1(t)+g_2(t)$ are considered with $g_1(t)=\theta g(t)$, $g_2(t)=(1-\theta)g(t)$ and $\theta\in[0,1]$. We thus obtain a system (1.1), (1.2) with

$$F_j(t,v) = A_jv + b_j(t) + g_j(t) \quad (j=1,2). \tag{2.3}$$

The errors of the LOD methods will be measured in the discrete L_2-norm

$$\|w\| = [h^2 \sum_{i,j=1}^{m} |w_{ij}|^2]^{1/2} \quad \text{for } w = (w_{ij}) \in \mathbb{R}^M.$$

The induced spectral norm for the space $L(\mathbb{R}^M)$ of $M \times M$ matrices is also denoted by $\|\cdot\|$. Along with the property that A_1 and A_2 commute we shall often use the fact that these matrices are negative definite. This implies

$$\|r(\tau A_j)\| \leq 1, \; \|(I - \tfrac{1}{2}\tau A_j)^{-1}\| \leq 1 \; \text{ and } \; \|(I - \tfrac{1}{2}\tau A_j)^{-1}\tau A_j\| \leq 2 \tag{2.4}$$

for all $\tau > 0$, $j = 1,2$, where $r(z) = (1 - \frac{1}{2}z)^{-1}(1 + \frac{1}{2}z)$ is the stability function of the implicit midpoint rule.

The fully discrete numerical solution u_n will be compared with the exact PDE solution $\mathbf{u}$. This exact solution is assumed, throughout all the paper, to be sufficiently often differentiable in time and space. The restriction of $\mathbf{u}(x,y,t)$ to Ω_h is denoted by $u_h(t)$, and it will be assumed that $u_0 = u_h(0)$. The error due to space discretization is described by

$$\alpha_h(t) = \dot{u}_h(t) - F(t, u_h(t)). \tag{2.5}$$

For our model problem this spatial error is $O(h^2)$.

The main objective of the present investigation is the temporal order of convergence p appearing in the error bound

$$\|u_h(t_n) - u_n\| \leq C\tau^p + Dh^2 \quad (\tau, h > 0, \; 0 \leq t_n \leq T) \tag{2.6}$$

with constants C, D independent of τ and h. Here τ and h are allowed to tend to zero simultaneously and independently of each other (unconditional convergence).

2.2. Some technical results

In the error bounds it will be required that the constants involved are not affected by τ and h. The symbol $O(\tau^p h^q)$ will be used to denote a scalar, or vector, whose absolute value, or L_2-norm, is bounded by $C\tau^p h^q$ for all possible $\tau, h > 0$ with $C > 0$ independent of τ and h. In particular, $O(\tau^p)$ thus stands for a term bounded by $C\tau^p$ for $\tau > 0$ uniformly for $h > 0$. Naturally we always have $\tau \leq T$ and $h \leq 1$.

Let $\otimes$ denote the left Kronecker (or tensor) product of vectors in $\mathbb{R}^m$ and matrices in $L(\mathbb{R}^m)$, i.e.,

$$v \otimes w = (w_1 v^T, w_2 v^T, \ldots, w_m v^T)^T \in \mathbb{R}^M \quad \text{for} \quad v, w = (w_i) \in \mathbb{R}^m$$

and $P \otimes S$ stands for the block matrix in $L(\mathbb{R}^M)$ with blocks $s_{ij}P \in L(\mathbb{R}^m)$ for $P, S = (s_{ij}) \in L(\mathbb{R}^m)$. Standard properties of such products can be found in [3], [7], for example. (We use the left Kronecker product rather than the more common right form since it gives a more natural notation here.) A gridfunction with values $v(x_i)w(y_j)$ on Ω_h can be written as vector $v\otimes w$ with $v_i=v(x_i), w_j=w(y_j)$. It is easy to verify that

$$\|v\otimes w\| = |v|\,|w| \quad \text{for } v,w\in\mathbb{R}^m$$

where $|\cdot|$ is the 1-dimensional discrete L_2-norm

$$|v| = [hv^Tv]^{1/2} \quad \text{for } v\in\mathbb{R}^m .$$

Let $Q=h^{-2}$ tridiag $(-1,2,-1)\in L(\mathbb{R}^m)$ be the standard 1-dimensional approximation for $-d^2/dx^2$. The matrices A_1, A_2 can be written as

$$A_1 = -Q\otimes I, \; A_2 = -I\otimes Q,$$

I being the identity matrix. The eigenvalues and eigenvectors of Q are given by

$$\begin{aligned} Q\phi_i &= \lambda_i\phi_i \; (1\leq i\leq m), \; \lambda_i=4h^{-2}\sin^2(ih\pi/2), \\ \phi_i &= \sqrt{2}(\sin(ih\pi), \sin(2ih\pi),\ldots,\sin(mih\pi))^T\in\mathbb{R}^m . \end{aligned} \tag{2.7}$$

Since $\{\phi_1,\phi_2,\ldots,\phi_m\}$ is an orthonormal set in $\mathbb{R}^m$ w.r.t. the inner product hv^Tw, we have for any $v\in\mathbb{R}^m$ the Fourier decomposition

$$v = \sum_{i=1}^m \hat{v}_i\phi_i, \;\; \hat{v}_i = hv^T\phi_i \;\; (1\leq i\leq m) \;\; \text{and} \;\; |v|^2 = \sum_{i=1}^m |\hat{v}_i|^2.$$

In the remainder of this section we consider a rational function ψ such that

$$\psi(0)=0 \;\; \text{and} \;\; |\psi(z)|\leq K \quad \text{for all } z\leq 0. \tag{2.8}$$

Since τQ is positive definite, it easily follows that for any $v\in\mathbb{R}^m$ a constant C exists such that

$$|\psi(-\tau Q)v|\leq C \quad \text{for all } \tau,h>0 \tag{2.9}$$

(we can take $C=K|v|$; for this inequality the assumption $\psi(0)=0$ is irrelevant). If $v\in\mathbb{R}^m$ can be viewed as a smooth 1-dimensional gridfunction a better estimate is known for small τ, also valid uniformly for $h>0$.

LEMMA 2.1. *Let $\chi\in C^2[0,1]$ and $v=(v_i)\in\mathbb{R}^m$ with $v_i=\chi(ih)$. Assume (2.8). Then, for any $\gamma<\frac{1}{4}$ there exists a $C_\gamma>0$ such that*

$$|\psi(-\tau Q)v|\leq C_\gamma\tau^\gamma \quad \textit{for all } \tau,h>0. \qquad \square$$

This lemma can be proved as in [6], Appendix, where $\psi(z)=(1-\frac{1}{2}z)^{-1}z$ was considered. The constant C_γ can be bounded in terms of γ and upper bounds for $|z^{-\gamma}\psi(z)|$ $(z\leq 0)$ and $|\chi(0)|$, $|\chi(1)|$, $|\chi''(x)|$ $(0\leq x\leq 1)$. The function χ thus may depend on h as long as these upper bounds are valid uniformly in h.

Sharpness of the above result can be shown by considering the vector $e=(1,1,...,1)^T \in \mathbb{R}^m$.

LEMMA 2.2. *Let $\beta>0, \psi \neq 0$ and assume (2.8). There exists a $C>0$ such that*

$$|\psi(-\tau Q)e| \geq C\tau^{1/4} \quad \textit{for } \tau,h>0 \textit{ with } \tau/h^2 \geq \beta. \qquad \square$$

Also this lemma can be proved as in [6], Appendix.

Note that if we were considering only a fixed meshwidth $h=h_0>0$, then for any $v \in \mathbb{R}^m$ it would hold that

$$|\psi(-\tau Q)v| \leq C_0 \tau \quad \text{for } \tau>0,\ h=h_0 \tag{2.10}$$

where $C_0=|Q|\ |v| \cdot \sup\{|z^{-1}\psi(z)|: z \leq 0\}$. However, since the spectral norm of Q equals $|\lambda_m| \approx 4h^{-2}$ such bounds are useless for small h. From the above estimates it is seen that bounds valid for all $h>0$ are possible, but at the cost of a lower power of τ. This is essentially the cause for order reduction if $h \downarrow 0$ in time integration methods.

For vectors $w \in \mathbb{R}^M$, which can be regarded as gridfunctions on the 2-dimensional grid Ω_h, we have

$$w = \sum_{i,j=1}^{m} \hat{w}_{ij} \phi_i \otimes \phi_j \ , \quad \|w\|^2 = \sum_{i,j=1}^{m} |\hat{w}_{ij}|^2$$

with Fourier components $\hat{w}_{ij}=h^2 w^T[\phi_i \otimes \phi_j]$. Lemma 2.1 yields the following 2-dimensional result.

LEMMA 2.3. *Let $\chi \in C^2([0,1]^2)$ and $w=(w_{ij}) \in \mathbb{R}^M$ with $w_{ij}=\chi(ih,jh)$. Assume (2.8). Then, for any $\gamma<1/4$ there is a $C_\gamma>0$ such that*

$$\|\psi(\tau A_j)w\| \leq C_\gamma \tau^\gamma \quad \textit{for all } \tau,h>0 \textit{ and } j=1,2.$$

PROOF. Consider $j=1$ and the row-wise ordering $w=(w_1^T,...,w_m^T)^T$, $w_j=(w_{1j},...,w_{mj})^T$. We have, cf. [7],

$$\psi(\tau A_1) = \psi(-\tau Q \otimes I) = \psi(-\tau Q) \otimes I$$

and consequently

$$\psi(\tau A_1)w = ((\psi(-\tau Q)w_1)^T,...,(\psi(-\tau Q)w_m)^T)^T.$$

The bound for $\|\psi(\tau A_1)w\|$ now easily follows from Lemma 2.1.

For $j=2$ we can proceed in the same way by using a columnwise ordering for w. $\square$

3. Recursions for the discretization errors

3.1. *The basic scheme*

In this section recursions for the global discretization errors $u_h(t_n)-u_n$ of the LOD schemes will be derived. As we shall see, the propagation of errors is governed by the matrix

$$R = r(\tau A_1)\, r(\tau A_2). \tag{3.1}$$

Since it holds, in view of (2.4), that $\|R\| \leq 1$, stability is always ensured. In this section most attention will be given to the structure of the temporal discretization errors introduced in one single step of the integration process. These locals discretization errors will be expressed in terms of derivatives of $u_h(t)$ and

$$v_h(t) = F_1(t,u_h(t)) - F_2(t,u_h(t)). \tag{3.2}$$

Note that although F_1 and F_2 contain negative powers of h, the function v_h and its derivatives are bounded uniformly for $h>0$ provided that PDE solution $\mathbf{u}$ is sufficiently smooth (which will always be assumed).

Consider along with (1.4) a perturbed scheme

$$\tilde{u}_{n+\frac{1}{2}} = \tilde{u}_n + \tau F_1(t_n+\tfrac{1}{4}\tau, \tfrac{1}{2}\tilde{u}_n + \tfrac{1}{2}\tilde{u}_{n+\frac{1}{2}} + p_n) + q_n, \tag{3.3a}$$

$$\tilde{u}_{n+1} = \tilde{u}_{n+\frac{1}{2}} + \tau F_2(t_n+\tfrac{3}{4}\tau, \tfrac{1}{2}\tilde{u}_{n+\frac{1}{2}} + \tfrac{1}{2}\tilde{u}_{n+1} + p_{n+\frac{1}{2}}) + q_{n+\frac{1}{2}}. \tag{3.3b}$$

The perturbations p_k, q_k may represent various error sources, for instance round off. Below these perturbations will be used to derive suitable expressions for the local discretizations errors.

Let $\epsilon_k = \tilde{u}_k - u_k$ for $k=n, n+\frac{1}{2}$ and $n \geq 0$. By subtracting (1.4) form (3.3) we obtain

$$\epsilon_{n+\frac{1}{2}} = \epsilon_n + \tfrac{1}{2}\tau A_1(\epsilon_n + \epsilon_{n+\frac{1}{2}}) + \tau A_1 p_n + q_n,$$

$$\epsilon_{n+1} = \epsilon_{n+\frac{1}{2}} + \tfrac{1}{2}\tau A_2(\epsilon_{n+\frac{1}{2}} + \epsilon_{n+1}) + \tau A_2 p_{n+\frac{1}{2}} + q_{n+\frac{1}{2}}.$$

Elimination of $\epsilon_{n+\frac{1}{2}}$ leads to

$$\epsilon_{n+1} = R\epsilon_n + \delta_n \quad (n=0,1,2,...) \tag{3.4}$$

where

$$\delta_n = r(\tau A_2)(I-\tfrac{1}{2}\tau A_1)^{-1}\,[\tau A_1 p_n + q_n] + (I-\tfrac{1}{2}\tau A_2)^{-1}\,[\tau A_2 p_{n+\frac{1}{2}} + q_{n+\frac{1}{2}}]. \tag{3.5}$$

If we put $\tilde{u}_n = u_h(t_n)$ for all $n \geq 0$, (3.4) gives a recursion for the global discretization errors. Apparently, δ_n is then a local discretization error. To obtain a suitable expression for these local errors, let $\tilde{u}_{n+\frac{1}{2}} = u_h(t_{n+\frac{1}{2}})$ and

$$p_n = -\frac{1}{2}u_h(t_n) + u_h(t_n + \frac{1}{4}\tau) - \frac{1}{2}u_h(t_n + \frac{1}{2}\tau),$$

$$p_{n+\frac{1}{2}} = -\frac{1}{2}u_h(t_n + \frac{1}{2}\tau) + u_h(t_n + \frac{3}{4}\tau) - \frac{1}{2}u_h(t_n + \tau).$$

This choice is made so that all perturbations p_k, q_k only depend on quantities like $u_h(t), F_j(t, u_h(t))$ and time derivatives thereof which are known be bounded for $h \downarrow 0$.

By a Taylor expansion around $t = t_{n+\frac{1}{2}}$ and using the relation $\dot{u}_h = F(t, u_h) + O(h^2)$ it follows with the above choice that

$$p_n = -\frac{1}{32}\tau^2 \ddot{u}_h(t_{n+\frac{1}{2}}) + O(\tau^3), \quad p_{n+\frac{1}{2}} = -\frac{1}{32}\tau^2 \ddot{u}_h(t_{n+\frac{1}{2}}) + O(\tau^3),$$

$$q_n = -\frac{1}{2}\tau v_h(t_{n+\frac{1}{2}}) + \frac{1}{8}\tau^2 \dot{v}_h(t_{n+\frac{1}{2}}) + O(\tau^3) + O(\tau h^2),$$

$$q_{n+\frac{1}{2}} = \frac{1}{2}\tau \; v_h(t_{n+\frac{1}{2}}) + \frac{1}{8}\tau^2 \dot{v}_h(t_{n+\frac{1}{2}}) + O(\tau^3) + O(\tau h^2),$$

with v_h given by (3.2). Relation (3.5) can be written as

$$\delta_n = (I - \frac{1}{2}\tau A_1)^{-1}(I - \frac{1}{2}\tau A_2)^{-1}[(I + \frac{1}{2}\tau A_2)(\tau A_1 p_n + q_n) + (I - \frac{1}{2}\tau A_1)(\tau A_2 p_{n+\frac{1}{2}} + q_{n+\frac{1}{2}})].$$

Using the bounds (2.4) we obtain after some calculations

$$\delta_n = (I - \frac{1}{2}\tau A_1)^{-1}(I - \frac{1}{2}\tau A_2)^{-1} \, [\frac{1}{4}\tau^2 \dot{v}_h(t_{n+\frac{1}{2}}) - \frac{1}{4}\tau^2 A \; v_h(t_{n+\frac{1}{2}}) + \tag{3.6}$$

$$+ \frac{1}{16}\tau^3 (A_2 - A_1)\dot{v}_h(t_{n+\frac{1}{2}}) - \frac{1}{32}\tau^3 A \ddot{u}_h(t_{n+\frac{1}{2}})] + O(\tau^3) + O(\tau h^2).$$

So, the global discretization errors $\epsilon_n = u_h(t_n) - u_n$ of the basic LOD scheme (1.5) satisfy recursion (3.4) with local errors δ_n given by (3.6). In the Sections 4 and 5 this will be used to derive error bounds.

3.2. The alternating schemes

For the alternating schemes (1.6), (1.7) recursions for the global errors $\epsilon_n = u_h(t_n) - u_n$ are easily obtained from the previous results. Let δ_n be the local errors, given by (3.6), of the basic scheme $u_{n+1} = \Phi_{1,2}(u_n)$. The scheme $u_{n+1} = \Phi_{2,1}(u_n)$ has local errors

$$\delta'_n = (I - \frac{1}{2}\tau A_1)^{-1}(I - \frac{1}{2}\tau A_2)^{-1} \, [-\frac{1}{4}\tau^2 \dot{v}_h(t_{n+\frac{1}{2}}) + \frac{1}{4}\tau^2 A v_h(t_{n+\frac{1}{2}}) + \tag{3.7}$$

$$+ \frac{1}{16}\tau^3 (A_2 - A_1) \, \dot{v}_h \, (t_{n+\frac{1}{2}}) - \frac{1}{32}\tau^3 A \ddot{u}_h(t_{n+\frac{1}{2}})] + O(\tau^3) + O(\tau h^2).$$

This follows directly from (3.2), (3.6) by interchanging the indices 1,2.

For the scheme (1.6) we thus have

$$\epsilon_{n+1} = R\epsilon_n + \delta_n \ , \ \epsilon_{n+2} = R\epsilon_{n+1} + \delta'_{n+1} \quad (n=0,2,4,...)\,. \tag{3.8}$$

Taking the two steps together it follows that

$$\epsilon_{n+2} = R^2\epsilon_n + \rho_n \ , \ \rho_n = R\delta_n + \delta'_{n+1} \quad (n=0,2,4,...). \tag{3.9}$$

Likewise, we get for the parallel alternating scheme (1.7) the error recursion

$$\epsilon_{n+1} = R\epsilon_n + \sigma_n \ , \ \sigma_n = \frac{1}{2}\delta_n + \frac{1}{2}\delta'_n \quad (n=0,1,2,...)\,. \tag{3.10}$$

This local error σ_n can be written as

$$\sigma_n = (I-\frac{1}{2}\tau A_1)^{-1}(I-\frac{1}{2}\tau A_2)^{-1}\,[\frac{1}{16}\tau^3(A_2-A_1)\dot{v}_h(t_{n+\frac{1}{2}})-\frac{1}{32}\tau^3 A\ddot{u}_h(t_{n+\frac{1}{2}})] + \tag{3.11}$$

$$+\ O(\tau^3) + O(\tau h^2).$$

4. Local error bounds

4.1. The basic scheme

For a time integration method consistent of order p one would expect local discretization errors to behave like $O(\tau^{p+1})+O(\tau h^2)$, with the τh^2 contribution due to spatial errors.

Consider the local error δ_n, given by (3.6), of the basic scheme (1.5) which has order 1 in the classical ODE sense. It is easily seen, by using (2.4), that $\|\delta_n\| = O(\tau^2)+O(\tau h^2)$ provided that

$$\|Av_h(t)\| = O(1) \tag{4.1}$$

uniformly in h. This condition, however, will not hold in general, due to the fact that even if the PDE solution $\mathbf{u}$ is very smooth, $v_h(t)$ (which then also is a smooth gridfunction) need not be zero near the boundaries Γ. For smooth $\mathbf{u}$, (4.1) is equivalent to the compatibility condition

$$(\frac{\partial}{\partial x^2} - \frac{\partial^2}{\partial y^2})\,\mathbf{u}\,(x,y,t) + (2\theta-1)\,\mathbf{g}\,(x,y,t) = 0 \quad \text{on } \Gamma, \tag{4.2}$$

and this will only be satisfied in exceptional cases, e.g. if $\mathbf{g}\equiv 0$, $\mathbf{u}_\Gamma \equiv 0$.

In the following it will not be assumed that (4.1) holds. It is then still possible to obtain upper bounds for the local errors uniformly for $h>0$, but with a reduced order of consistency.

Lemma 4.1. *For any $\gamma<1/4$ there are constants $C_\gamma, D>0$ such that*

$$\|\delta_n\| \leq C_\gamma \tau^{1+\gamma} + D\tau h^2 \quad \textit{for all } \tau,h>0,\ 0\leq t_n \leq T.$$

PROOF. From (3.6) and the bounds (2.4) it follows that

$$\delta_n = -\frac{1}{4}\tau^2 (I-\frac{1}{2}\tau A_1)^{-1}(I-\frac{1}{2}\tau A_2)^{-1} A v_h(t_{n+\frac{1}{2}}) + O(\tau^2) + O(\tau h^2)$$

and

$$\|\delta_n\| \leq \frac{1}{4}\tau^2 \sum_{j=1}^{2} \|(I-\frac{1}{2}A_j)^{-1} A_j v_h(t_{n+\frac{1}{2}})\| + O(\tau^2) + O(\tau h^2).$$

Since $v_h(t)$ is a smooth gridfunction, provided **u** is smooth, the result now follows from Lemma 2.3.

□

In the above estimate we have obtained a temporal order of consistency ≈¼ only. To show that this result in sharp we will consider problem (2.1) with a suitably chosen simple solution. This solution can be taken stationary.

EXAMPLE 4.2. Consider distribution of the source term with $\theta=0$; other natural choices $\theta=\frac{1}{2}$ or $\theta=1$ can be treated similarly. Let

$$\mathbf{u}(x,y,t) = \frac{1}{2}x(1-x)\sin(\pi y). \tag{4.3}$$

Then $g_2(t)=g(t)$ is the restriction of $-\Delta\mathbf{u}(x,y,t)$ to Ω_h, and $b(t)=0$ since **u** satisfies homogeneous boundary conditions. It follows by some simple calculations that

$$v_h(t) = -2e\otimes\phi_1 + O(h^2) \tag{4.4}$$

where $e=(1,1,...,1)^T\in\mathbf{R}^m$ and ϕ_1 is the first eigenvector of Q, cf. Section 2.2.

To simplify the situation further, we neglect in the following the spatial errors $O(h^2)$; it is only the temporal order which is of interest here. So, we consider the local error

$$\delta_n = \frac{1}{2}\tau^2 (I-\frac{1}{2}\tau A_1)^{-1}(I-\frac{1}{2}\tau A_2)^{-1} A\,[e\otimes\phi_1]. \tag{4.5}$$

This can be written as

$$\delta_n = -\frac{1}{2}\tau^2\,[(I+\frac{1}{2}\tau Q)^{-1} Q e\otimes(I+\frac{1}{2}\tau Q)^{-1}\phi_1] - \tag{4.6}$$

$$-\frac{1}{2}\tau^2[(I+\frac{1}{2}\tau Q)^{-1} e\otimes(I+\frac{1}{2}\tau Q)^{-1} Q\phi_1].$$

The second term on the righthand ride is of $O(\tau^2)$ since $Q\phi_1=\lambda_1\phi_1=O(1)$ and $(I+\frac{1}{2}\tau Q)^{-1}$ is also bounded for all $h>0$. Using Lemma 2.2 it can be seen that the norm of the first term can be bounded from below by $C\tau^{5/4}$, $C>0$, provided that $\tau/h^2 \neq o(1)$, and consequently such a lower bound also holds for $\|\delta_n\|$. Hence the order result in Lemma 4.1 cannot be essentially improved if both τ and h tend to zero and $\tau/h^2 \geq \beta > 0$. □

4.2. The sequentially alternating scheme

Consider the local error ρ_n in (3.9) of scheme (1.6). By some calculations it can be seen that the temporal order of consistency is 2 provided **u** satisfies certain compatibility conditions or if $h \geq h_0 > 0$. So, it seems that compared to the basic scheme (1.5) one order is gained. Below it will be shown however that this gain is lost in general if $h \downarrow 0$.

Observe that, since $\|R\| \leq 1$,

$$\|\rho_n\| \leq \|\delta_n\| + \|\delta'_{n+1}\|.$$

The bound of Lemma 4.1 for $\|\delta_n\|$ holds for $\|\delta'_{n+1}\|$ as well. Therefore the following result is directly obtained.

LEMMA 4.3. *For any $\gamma < 1/4$ there are $C_\gamma, D > 0$ such that*

$$\|\rho_n\| \leq C_\gamma \tau^{1+\gamma} + D\tau h^2 \quad \textit{for all } \tau, h > 0, \ 0 \leq t_n \leq T. \qquad \square$$

To prove sharpness we consider again a simple stationary problem.

EXAMPLE 4.4. Let, as in Example 4.2, **u** be given by (4.3) and $\theta = 0$. Omitting the spatial errors, it follows from (3.7), (4.4), (4.5) that $\delta'_{n+1} = -\delta_n$ and

$$\rho_n = (R - I)\delta_n = \frac{1}{2}\tau^3 (I - \frac{1}{2}\tau A_1)^{-2} (I - \frac{1}{2}\tau A_2)^{-2} A^2 \, [e \otimes \phi_1]\,. \tag{4.7}$$

It follows that

$$\rho_n = \frac{1}{2}\tau^3 [(I + \frac{1}{2}\tau Q)^{-2} Q^2 e \otimes (I + \frac{1}{2}\tau Q)^{-2} \phi_1] +$$

$$+ \tau^3 [(I + \frac{1}{2}\tau Q)^{-2} Q e \otimes (I + \frac{1}{2}\tau Q)^{-2} Q \phi_1] + \frac{1}{2}\tau^3 [(I + \frac{1}{2}\tau Q)^{-2} e \otimes (I + \frac{1}{2}\tau Q)^{-2} Q^2 \phi_1].$$

Now, let $\tau, h \downarrow 0$ and $\tau/h^2 \geq \beta > 0$. Using Lemma 2.2 it is seen that norm of the first term on the right-hand side is bounded from below by $C\tau^{5/4}, C > 0$. The second and third term are $O(\tau^{9/4}), O(\tau^3)$, respectively. Hence for $\|\rho_n\|$ we have a lower bound $C'\tau^{5/4}$ with a constant $C' > 0$. □

4.3. The parallel alternating scheme

The local error σ_n of the parallel alternating schem (1.7) is given (3.11). In the same way as in Lemma 4.1 it can be shown that following local bound is valid.

LEMMA 4.5. *For any $\gamma < 1/4$ there are $C_\gamma, D > 0$ such that*

$$\|\sigma_n\| \leq C_\gamma \tau^{2+\gamma} + D\tau h^2 \quad \textit{for all } \tau, h > 0,\ 0 \leq t_n \leq T. \qquad \square$$

Compared to the other two schemes one order of τ has been gained. Note however that there still is an order reduction, since for fixed ODEs scheme (1.7) is consistent with order 2. To show that the above bound is sharp when $h \downarrow 0$ we consider again a simple, constructed example.

EXAMPLE 4.6. Suppose $\theta = 0$. As before, other choices for θ can be treated similarly. Let

$$\mathbf{u}(x,y,t) = \frac{1}{2} t x (1-x) \ \sin(\pi y).$$

Note that here a nonstationary solution is considered; for stationary **u** the scheme (1.7) is exact in time. For the above problem we have $\ddot{u}_h(t) = 0$ and

$$\dot{v}_h(t) = -2e \otimes \phi_1 + O(h^2).$$

Hence, omitting the spatial errors,

$$\sigma_n = -\frac{1}{8}\tau^3 (I - \frac{1}{2}\tau A_1)^{-1} (I - \frac{1}{2}\tau A_2)^{-1} (A_2 - A_1)\, [e \otimes \phi_1],$$

which can be written as

$$\sigma_n = -\frac{1}{8}\tau^3 [(I + \frac{1}{2}\tau Q)^{-1} Qe \otimes (I + \frac{1}{2}\tau Q)^{-1} \phi_1] + $$
$$+ \frac{1}{8}\tau^3 [(I + \frac{1}{2}\tau Q)^{-1} e \otimes (I + \frac{1}{2}\tau Q)^{-1} Q\phi_1].$$

Comparing this with formula (4.6), it now follows that $\|\sigma_n\| \geq C\tau^{9/4}$ with $C > 0$, if $\tau, h \downarrow 0$ and $\tau / h^2 \geq \beta > 0$. $\square$

In the examples treated in this section the order reduction has been caused by the presence of a source term **g**. The same effect would have been obtained if we had taken $\mathbf{g} \equiv 0$ but boundary values $\mathbf{u}_\Gamma$ varying in time , for instance with

$$\mathbf{u}(x,y,t) = \exp(-\frac{5}{4}\pi^2 t) \ \sin(\pi x) \ \sin(\pi y / 2).$$

5. GLOBAL ERROR BOUNDS

5.1. Global order reduction

Consider an error recursion in $\mathbb{R}^M$ of the form

$$e_{n+1} = Se_n + d_n \quad (n = 0, 1, \ldots, N-1),\ e_0 = 0 \tag{5.1}$$

where $N = \lfloor T/\tau \rfloor$ and $S \in L(\mathbb{R}^M)$, $\|S\| \leq 1$. The vectors e_n and d_n stand for global and local errors, respectively. If $\|d_n\| = O(\tau^{1+\gamma}) + O(\tau h^2)$ uniformly for $n \geq 0$, we obtain in the standard way the global result $\|e_n\| = O(\tau^\gamma) + O(h^2)$, i.e., convergence with temporal order γ. In many cases however it is possible to improve this global result by taking into account certain cancellation effects, giving a temporal order of convergence which is larger than the temporal order of consistency. Consequently the order reduction in the global errors is often less prominent than in the local errors; it may even be absent. In this section this will be shown to be the case for the schemes (1.5) and (1.7) (not for scheme (1.6) though). Similar results for Runge-Kutta methods can be found in [2], [11], for example, and for an ADI method in [6].

Let $\alpha > 0$ and consider the following two statements with constants $C, C' > 0$,

$$\|e_n\| \leq C\tau^\alpha \quad (0 \leq n \leq N), \tag{5.2}$$

$$d_n = (I-S)\xi_n + \eta_n \ (0 \leq n \leq N-1) \ \text{with} \ \xi_n, \eta_n \in \mathbb{R}^M \quad \text{such that} \tag{5.3}$$

$$\|\xi_n\| \leq C'\tau^\alpha, \|\eta_n\| \leq C'\tau^{1+\alpha} (0 \leq n \leq N-1) \ \text{and} \ \|\xi_n - \xi_{n-1}\| \leq C'\tau^{1+\alpha} \ (1 \leq n \leq N-1).$$

Lemma 5.1. *Assume (5.3). Then (5.2) holds with C only depending on C′ and T.*

Proof. (cf. also [5]). Instead of directly bounding the local errors d_n, we first elaborate the recursion (5.1), giving

$$e_n = d_{n-1} + Sd_{n-2} + \cdots + S^{n-1}d_0 =$$
$$= \xi_{n-1} - S(\xi_{n-1} - \xi_{n-2}) - \cdots - S^{n-1}(\xi_1 - \xi_0) - S^n\xi_0 +$$
$$+ \eta_{n-1} + S\eta_{n-2} + \cdots + S^{n-1}\eta_0.$$

Hence, by bounding the terms in this last sum,

$$\|e_n\| \leq 2C'\tau^\alpha + 2nC'\tau^{1+\alpha} \leq 2(1+T)C'\tau^\alpha \quad \text{for } n\tau \leq T.$$ □

In case the local errors are constant the reverse implication also holds.

Lemma 5.2. *Assume (5.2) and* $d_n = d_0$ $(0 \leq n \leq N-1)$. *Then there are* $\xi_n = \xi_0, \eta_n = \eta_0$ *such that (5.3) holds, with C′ only depending on C and T.*

Proof. Let $\lambda = (1+\tau)^{-1}$, so that $\|\lambda S\| < 1$. Consider a sequence $\{v_n\}$ defined by

$$v_{n+1} = \lambda S v_n + d_0 \ (n \geq 0), \ v_0 = 0.$$

Then, for $w_n = v_n - e_n$, we have,

$$w_{n+1} = \lambda S w_n - (1-\lambda)S\, e_n \ (n \geq 0), \ w_0 = 0,$$

and hence

$$\|w_{n+1}\| \leq \lambda \|w_n\| + (1-\lambda) C\tau^{\alpha} \quad (n \geq 0), \ \|w_0\| = 0.$$

This leads in a standard way to the global result

$$\|w_n\| \leq (1+\lambda+ \cdots +\lambda^{n-1})\,(1-\lambda)C\tau^{\alpha} = (1-\lambda^n)C\tau^{\alpha} \leq C\tau^{\alpha}.$$

Thus we obtain

$$\|v_n\| \leq \|e_n\| + \|w_n\| \leq 2C\tau^{\alpha} \quad (n \geq 0).$$

On the other hand, from the recursion for the v_n it follows that

$$v_N = (I+\lambda S+ \cdots +\lambda^{N-1}S^{N-1})d_0 = (I-\gamma^N S^N)(I-\lambda S)^{-1} d_0.$$

Therefore

$$\|(I-\lambda S)^{-1} d_0\| \leq (1-\lambda^N)^{-1} 2C\tau^{\alpha} \leq C'\tau^{\alpha}$$

with $C' = 2C(1+T)/T$.

Now take

$$\xi_0 = (I-\lambda S)^{-1} d_0 \ , \ \eta_0 = (1-\lambda)S\xi_0.$$

Then it easily seen that $d_0 = (I-S)\xi_0 + \eta_0$ and from the above it follows that $\|\xi_0\| \leq C'\tau^{\alpha}$, $\|\eta_0\| \leq C'\tau^{1+\alpha}$. □

It should be noted that if $S = I + O(\tau)$, then (5.3) is just equivalent with $\|d_n\| = O(\tau^{1+\alpha})$, and (5.2) would follow in the standard way. The decomposition (5.3) can be useful if S is bounded away from I, as will be the case if $S = R = r(\tau A_1)r(\tau A_2)$ when $h \downarrow 0$. Then (5.3) merely implies $\|d_n\| = O(\tau^{\alpha})$.

REMARK 5.3. For simplicity, only temporal errors were considered in the above, but it easily follows that

$$\|e_n\| \leq C\tau^{\alpha} + Dh^2 \quad (0 \leq n \leq N)$$

if (5.3) is satisfied with $\|\xi_n\| \leq C'\tau^{\alpha}, \|\xi_n - \xi_{n-1}\| \leq C'\tau^{1+\alpha}$ and $\|\eta_n\| \leq C'\tau^{1+\alpha} + D'\tau h^2$. □

5.2. *The basic scheme*

Scheme (1.5) is convergent with temporal order 1 in the classical ODE sense. The following theorem shows that this remains valid as $h \downarrow 0$. So, in the transition from local to global error the order reduction disappears.

THEOREM 5.4. *Consider (1.5) and (2.1). There are constants $C, D > 0$, depending only on T and the* *smoothness of* **u**, *such that*

$$\|u_h(t_n)-u_n\|\leq C\tau+Dh^2 \quad \textit{for all } \tau,h>0,\ 0\leq t_n\leq T.$$

PROOF. Let

$$\xi_n = \frac{1}{4}\tau v_h(t_{n+\frac{1}{2}}),\ \eta_n = \delta_n-(I-R)\xi_n.$$

Smoothness of **u** implies $\|\xi_n\|=O(\tau), \|\xi_n-\xi_{n-1}\|=O(\tau^2)$. By observing that

$$I-R=-\tau(I-\frac{1}{2}\tau A_2)^{-1}(I-\frac{1}{2}\tau A_2)^1 A, \tag{5.4}$$

it can be seen from (2.4) that $\|\eta_n\|=O(\tau^2)+O(\tau h)$. The proof follows from Lemma 5.1 and Remark 5.3. □

5.3. The sequentially alternating scheme

As we saw in Section 4.2, modifying (1.5) into (1.6) did not help to improve the temporal order of consistency. Here it will be shown that this modification even has a *negative* effect on the order of convergence in time.

EXAMPLE 5.5. Consider the problem of the Examples 4.2, 4.3 with $\tau=h$. Omitting the space errors we have for this stationary example the error recursion

$$\epsilon_{n+2} = R^2\epsilon_n + \frac{1}{2}\tau(R-I)^2\, e\otimes\phi_1\ (n=0,2,4,...)\,, \tag{5.5}$$

cf. (3.9), (4.7) and (5.4). According to Lemma 5.2 there will be convergence with temporal order α iff there are $\xi,\eta\in\mathbf{R}^M$ such that

$$\|\xi\| = O(\tau^\alpha)\,,\ \|\eta\| = O(\tau^{1+\alpha}), \tag{5.6a}$$

$$\frac{1}{2}\,\tau(R-I)^2\, e\otimes\phi_1 = (R^2-I)\xi+\eta. \tag{5.6b}$$

It will be shown that this necessarily implies $\alpha\leq½$.

We consider the Fourier expansions

$$\xi = \sum_{i,j=1}^{m}\hat{\xi}_{ij}\phi_i\otimes\phi_j\,,\ \eta=\sum_{i,j=1}^{m}\hat{\eta}_{ij}\phi_i\otimes\phi_j \text{ and } e = \sum_{i=1}^{m}\hat{e}_i\phi_i.$$

Let further $\lambda_1,...,\lambda_m$ be the eigenvalues of Q and $r_i=r(-\tau\lambda_i)$, cf. Section 2.2. By observing that $R\phi_i\otimes\phi_j=r_ir_j\phi_i\otimes\phi_j$, it follows from (5.6b) that

$$0 = (r_i^2r_j^2-1)\hat{\xi}_{ij} + \hat{\eta}_{ij} \text{ for } 1\leq i\leq m,\ 2\leq j\leq m.$$

This is fulfilled by taking $\hat{\xi}_{ij}=\hat{\eta}_{ij}=0$ for $j\geq 2$; any other choice would lead only to larger norms $\|\xi\|$ and $\|\eta\|$. With $j=1$ we obtain from (5.6b)

$$\frac{1}{2}\tau(r_1 r_i-1)^2\hat{e}_i = (r_1 r_i^2-1)\hat{\xi}_{1i} + \hat{\eta}_{1i} \text{ for } 1\leq i\leq m.$$

To simplify this relation, define

$$\hat{\xi}_i = 2(\frac{r_i-1}{r_1 r_i-1})^2 (\frac{r_1^2 r_i^2-1}{r_i^2-1})\hat{\xi}_{i1} \ , \ \hat{\eta}_i = 2(\frac{r_i-1}{r_1 r_i-1})^2\hat{\eta}_{i1}.$$

Since $r_1=1+O(\tau)$ it now follows that (5.6) reduces to

$$[\sum_{i=1}^{m}|\hat{\xi}_i|^2]^{½} = O(\tau^{\alpha}) \, , \, [\sum_{i=1}^{m}|\hat{\eta}_i|^2]^{½} = O(\tau^{1+\alpha}), \tag{5.7a}$$

$$\tau(r_i-1)^2\hat{e}_i = (r_i^2-1)\hat{\xi}_i + \hat{\eta}_i \text{ for } 1\leq i\leq m. \tag{5.7b}$$

In the following we will use the notation $f(x)\sim g(x)$ $(x\downarrow 0)$ for real functions f,g if there are $C_0,C_1,H>0$ such that $C_0 g(x)\leq f(x)\leq C_1 g(x)$ for $0<x<H$. From (2.7) it is easily seen that $\lambda_i\sim i^2$ $(h\downarrow 0)$ uniformly for $1\leq i\leq m$.

Now, consider the indices $i\geq\frac{1}{2}m$. As it is assumed that $\tau=h$, we have uniformly for these large indices

$$r_i-1\sim -1 \text{ and } r_i^2-1\sim -(r_i+1)\sim -\tau \ (\tau\downarrow 0).$$

Hence (5.7b) implies

$$\tau^2 \sum_{i\geq m/2}|\hat{e}_i|^2 \sim \sum_{i\geq m/2}|\hat{\eta}_i-\tau\hat{\xi}_i|^2. \tag{5.8}$$

From the proof of Lemma 2, p. 99 in [6] with $\gamma=0$ it can be seen that

$$\hat{e}_i = 0 \text{ for } i \text{ even} \, , \, \hat{e}_i=h\sqrt{2}\,\mathrm{cotg}(ih\pi/2) \text{ for } i \text{ odd},$$

and for $\tau=h\downarrow 0$

$$\sum_{i\geq m/2}|\hat{e}_i|^2 \sim h^2\sum_{x_i\geq ½}\mathrm{cotg}^2(x_i\pi/2) \sim h\int_{½}^{1}\mathrm{cotg}^2(x\pi/2)dx \sim h=\tau.$$

On the other hand we have in view of (5.7a)

$$\sum_{i\geq m/2}|\hat{\eta}_i-\tau\hat{\xi}_i|^2 = O(\tau^{2+2\alpha}).$$

Hence (5.8) implies $\tau^3=O(\tau^{2+2\alpha})$, and thus we must have $\alpha\leq ½$. □

Note that from the local error bound $\|\rho_n\|=O(\tau^{1+\gamma})+O(\tau h^2)$ of Lemma 4.3 with $\gamma\approx ¼$, it directly follows that the temporal order of convergence is always at least ¼, approximately. With the above example we now know that this order result is nearly optimal. (The question whether the order is ¼ or ½, or in between, is not so relevant since the convergence behaviour is very disappointing any way.)

So, we have the surprising result that scheme (1.6), which has a higher order in the classical ODE sense than (1.5), has a lower order of convergence in time when τ and h tend to 0 simultaneously. To give an illustration of this we present some numerical results for (2.1) with $T=2$ and

$$\mathbf{u}(x,y,t) = x(1-x)\,y\,(1-y)(16+y)$$

This solution is chosen such that no space errors are present and its magnitude is near 1. The source term $\mathbf{g}$ in (2.1) equals $-\Delta\mathbf{u}$, and we take $g_1=g_2=\frac{1}{2}g$ (i.e., $\theta=\frac{1}{2}$). The following Tables 5.1, 5.2 nicely illustrate the theory. On fixed space grids, where we are in the standard ODE situation, the sequentially alternating scheme (1.6) becomes more accurate than the basic scheme (1.5) for decreasing τ, but if both τ and h tend to 0 scheme (1.5) is the better one.

τ^{-1}	10	20	40	80	160
(1.5)	$0.47\,E-1$	$0.23\,E-1$	$0.12\,E-1$	$0.59\,E-2$	$0.29\,E-2$
(1.6)	$0.69\,E-1$	$0.32\,E-1$	$0.10\,E-1$	$0.27\,E-2$	$0.69\,E-3$

TABLE 5.1 Global errors (L_2-norm) for (1.5), (1.6) on a fixed space grid $h=1/5$.

τ^{-1}	10	20	40	80	160
(1.5)	$0.47\,E-1$	$0.34\,E-1$	$0.20\,E-1$	$0.10\,E-1$	$0.53\,E-2$
(1.6)	$0.69\,E-1$	$0.63\,E-1$	$0.45\,E-1$	$0.31\,E-1$	$0.22\,E-1$

TABLE 5.2. Global errors (L_2-norm) for (1.5), (1.6) with $h=2\tau$.

Although the theory has been formulated for the discrete L_2-norm, it is illuminating to consider the errors of the last table also in the maximum norm $\|v\|_\infty=\max|v_{ij}|$, $v=(v_{ij})\in\mathbb{R}^M$. In this norm the behaviour of the sequentially alternating scheme is even worse: apparently there is no convergence at all.

τ^{-1}	10	20	40	80	160
(1.5)	$0.74\,E-1$	$0.73\,E-1$	$0.46\,E-1$	$0.25\,E-1$	$0.13\,E-1$
(1.6)	0.11	0.14	0.14	0.14	0.14

TABLE 5.3. Global errors (max-norm) for (1.5), (1,6) with $h=2\tau$.

The fact that in the above tables the asymptotic behaviour does not show up for the larger τ values is probably caused by damping effects of terms $(I-\frac{1}{2}\tau A_j)^{-1}$, which will be stronger the larger τ is. In the error estimates we have merely used the bound $\|(I-\frac{1}{2}\tau A_j)^{-1}\|\leq 1$ for all $\tau>0$.

5.4. The parallel alternating scheme

As we saw in Section 4, the local errors of the parallel alternating scheme (1.7) also suffer from order reduction, though not as much as for scheme (1.6). Moreover, the following theorem shows that this reduction will be annihilated for the global errors of (1.7).

THEOREM 5.6. *Consider (1.7) and (2.1). There are constant $C,D>0$, depending only on T and the smoothness of* **u**, *such that*

$$\|u_h(t_n)-u_n\|\leq C\tau^2 + Dh^2 \quad \textit{for all } \tau,h>0,\ 0\leq t_n\leq T.$$

PROOF. Let

$$\xi_n = -\frac{1}{16}\tau^2 A^{-1}(A_2-A_1)\,\dot{v}_h(t_{n+\frac{1}{2}})+\frac{1}{32}\tau^2\ddot{u}_h(t_{n+\frac{1}{2}})\,,\ \eta_n = \sigma_n-(I-R)\xi_n.$$

Note that $A^{-1}(A_2-A_1)$ is symmetric with eigenvalues $(\lambda_i+\lambda_j)^{-1}(\lambda_i-\lambda_j)$ bounded by 1 in modulus, and hence the norm of this matrix is ≤ 1. Smoothness of **u** thus implies $\|\xi_n\|=O(\tau^2)$, $\|\xi_n-\xi_{n-1}\|=O(\tau^3)$. Further it is easily seen from (3.11), (5.4) that $\|\eta_n\|=O(\tau^3)+O(\tau h^2)$, and so the convergence result follows from Lemma 5.1 and Remark 5.3. □

In conclusion it can be said that the parallel alternating scheme (1.7) does what it is expected to do: its temporal order of convergence is 2, which is one higher than for the basic scheme. Moreover, the scheme (1.7) is exact in time for stationary solutions (so, in the Tables 5.1-5.3 all errors would be 0 for (1.7)).

6. SOME ALTERNATIVE FORMS

The choice of the time levels used in (1.4) seems somewhat arbitrary. We can consider the more general formula, with parameter c,

$$u_{n+\frac{1}{2}} = u_n + \tau F_1(t_n+c\tau,\frac{1}{2}u_n+\frac{1}{2}u_{n+\frac{1}{2}}), \tag{6.1a}$$

$$u_{n+1} = u_{n+\frac{1}{2}} + \tau F_2(t_n+(1-c)\tau,\ \frac{1}{2}u_{n+\frac{1}{2}}+\frac{1}{2}u_{n+1}). \tag{6.1b}$$

When using (6.1) as a basic scheme (1.5), the alternating schemes (1.6), (1.7) are of 2-nd order in the classical ODE sense for any choice of c. Apart from $c=¼$, which was used in (1.4), the choices $c=0$

or $c=½$ are also natural ones. Below it will be shown however that taking $c\neq¼$ will cause a global order reduction for the parallel alternating scheme if the boundary values $\mathbf{u}_\Gamma$ are not constant in time.

In the same way as in Section 3, by using suitable perturbations p_k, q_k, it can be shown that (6.1) has local errors

$$\delta_n = (I-\tfrac{1}{2}\tau A_1)^{-1}(I-\tfrac{1}{2}\tau A_2)^{-1}[(\tfrac{1}{2}-c)\tau^2\dot{v}_h(t_{n+\frac{1}{2}})-\tfrac{1}{4}\tau^2 A v_h(t_{n+\frac{1}{2}}) +$$

$$+ (c-\tfrac{1}{4})\tau^2(A_1-A_2)\dot{u}_h(t_{n+\frac{1}{2}})+(c-\tfrac{1}{4})\tau^3 A_1 A_2\dot{u}_h(t_{n+\frac{1}{2}}) +$$

$$+ (\tfrac{1}{4}c-\tfrac{1}{8})\tau^3(A_1-A_2)\dot{v}_h(t_{n+\frac{1}{2}}) +(\tfrac{1}{2}c^2-\tfrac{3}{4}c+\tfrac{1}{8})\tau^3 A\ddot{u}_h(t_{n+\frac{1}{2}})] + O(\tau^3) + O(\tau h^2).$$

The local error of the parallel alternating scheme (1.7), with (6.1) as basic scheme, can now seen to be

$$\sigma_n = (I-\tfrac{1}{2}\tau A_1)^{-1}(I-\tfrac{1}{2}\tau A_2)^{-1}[(c-\tfrac{1}{4})\tau^3 A_1 A_2\dot{u}_h(t_{n+\frac{1}{2}})+(\tfrac{1}{4}c-\tfrac{1}{8})\tau^3(A_1-A_2)\dot{v}_h(t_{n+\frac{1}{2}}) +$$

$$+ (\tfrac{1}{2}c^2-\tfrac{3}{4}c+\tfrac{1}{8})\tau^3 A\ddot{u}_h(t_{n+\frac{1}{2}})] + O(\tau^3) + O(\tau h^2).$$

In case the boundary conditions for $\mathbf{u}$ are constant in time it follows that $\dot{\mathbf{u}}=0$ on Γ and $\tau^3 A_1 A_2\dot{u}_h(t)=O(\tau^3)$. The term involving $\tau^3 A_1 A_2\dot{u}_h(t_{n+\frac{1}{2}})$ will lead however to a lower order of temporal convergence for time dependent boundary conditions.

EXAMPLE 6.1. Consider (2.1) with $\mathbf{g}\equiv 1$ and solution

$$\mathbf{u}(x,y,t) = t$$

Then $\dot{u}_h(t) = e\otimes e$ and $v_h(t)=(2\theta-1)e\otimes e$. Hence

$$\sigma_n = (c-\tfrac{1}{4})\tau^3(I-\tfrac{1}{2}\tau A_1)^{-1}(I-\tfrac{1}{2}\tau A_2)^{-1}A_1A_2[e\otimes e].$$

This can be written as

$$\sigma_n = (c-\tfrac{1}{4})\tau^3[(I+\tfrac{1}{2}\tau Q)^{-1}Qe\otimes(I+\tfrac{1}{2}\tau Q)^{-1}Qe].$$

Consequently

$$\|\sigma_n\| = |c-\tfrac{1}{4}|\tau^3|(I+\tfrac{1}{2}\tau Q)^{-1}Qe|^2 \geqslant |c-\tfrac{1}{4}|C^2\tau^{3/2}$$

for some $C>0$, provided that $\tau/h^2\neq o(1)$ (see Lemma 2.2).

Since σ_n is the error introduced in one step, it is clear that the order of convergence is $\leqslant 3/2$ if $c\neq¼$. In fact, it can be shown that there will be convergence with order 1 exactly when $\tau=h$. We shall not prove this result since it is of little relevance here. The important thing is simply that time dependent boundary conditions should be treated in (6.1) with $c=\frac{1}{4}$. This choice was also used in [13]. □

One can also derive LOD methods starting from the trapezoidal rule

$$u_{n+1} = u_n + \frac{1}{2}\tau F(t_n, u_n) + \frac{1}{2}\tau F(t_{n+1}, u_{n+1}) \tag{6.2}$$

instead of the implicit midpoint rule (1.3). An LOD method of this type is given by

$$u_{n+\frac{1}{2}} = u_n + \frac{1}{2}\tau F_1(t_n, u_n) + \frac{1}{2}\tau F_1(t_n + \frac{1}{2}\tau, u_{n+\frac{1}{2}}), \tag{6.3a}$$

$$u_{n+1} = u_{n+\frac{1}{2}} + \frac{1}{2}\tau F_2(t_n + \frac{1}{2}\tau, u_{n+\frac{1}{2}}) + \frac{1}{2}\tau F_2(t_{n+1}, u_{n+1}). \tag{6.3b}$$

For strongly nonlinear problems scheme (1.4) might possess better stability properties (cf. [14]), but for linear problems with constant coefficients (6.3) will also be stable.

An error recursion $\epsilon_{n+1} = R\epsilon_n + \delta_n$, $\epsilon_n = u_h(t_n) - u_n$, can be obtained by considering along with (6.3) a perturbed version

$$\tilde{u}_{n+\frac{1}{2}} = \tilde{u}_n + \frac{1}{2}\tau F_1(t_n, \tilde{u}_n) + \frac{1}{2}\tau F_1(t_n + \frac{1}{2}\tau, \tilde{u}_{n+\frac{1}{2}}) + q_n,$$

$$\tilde{u}_{n+1} = \tilde{u}_{n+\frac{1}{2}} + \frac{1}{2}\tau F_2(t_n + \frac{1}{2}\tau, \tilde{u}_{n+\frac{1}{2}}) + \frac{1}{2}\tau F_2(t_{n+1}, \tilde{u}_{n+1}) + q_{n+\frac{1}{2}}.$$

By taking $\tilde{u}_k = u_h(t_k)$ for $k = n, n+\frac{1}{2}$ and $n \geq 0$, we get

$$q_n = -\frac{1}{2}\tau v_h(t_{n+\frac{1}{2}}) + \frac{1}{8}\tau^2 \dot{v}_h(t_{n+\frac{1}{2}}) + O(\tau^3) + O(\tau h^2),$$

$$q_{n+\frac{1}{2}} = \frac{1}{2}\tau v_h(t_{n+\frac{1}{2}}) + \frac{1}{8}\tau^2 \dot{v}_h(t_{n+\frac{1}{2}}) + O(\tau^3) + O(\tau h^2).$$

In the same way as in Section 3, see the formulas (3.4), (3.5), it follows that the local discretization error for scheme (6.3) is given by

$$\delta_n = (I - \frac{1}{2}\tau A_1)^{-1}(I - \frac{1}{2}\tau A_2)^{-1} [\frac{1}{4}\tau^2 \dot{v}_h(t_{n+\frac{1}{2}}) - \frac{1}{4}\tau^2 A v_h(t_{n+\frac{1}{2}}) + \tag{6.4}$$

$$+ \frac{1}{16}\tau^3 (A_2 - A_1)\dot{v}_h(t_{n+\frac{1}{2}})] + O(\tau^3) + O(\tau h^2).$$

For the parallel alternating scheme based on (6.3) we get a local error

$$\sigma_n = (I - \frac{1}{2}\tau A_1)^{-1}(I - \frac{1}{2}\tau A_2)^{-1} [\frac{1}{16}\tau^3 (A_2 - A_1)\dot{v}_h(t_{n+\frac{1}{2}})] + O(\tau^3) + O(\tau h^2). \tag{6.5}$$

It can be seen, by following the previous proofs, that all error bounds for (1.5)-(1.7) remain unchanged if (6.3) is used as basic scheme. The error structure even is somewhat simpler with (6.3).

7. Concluding remarks

Of the three schemes (1.5), (1.6) and (1.7) considered in this paper, it is clear that the sequentially alternating scheme (1.6) is unsuited in its present form (i.e., without boundary corrections). The basic scheme (1.5) is first order accurate in time. For some practical problems 1-th order accuracy is sufficient, but even in such a situation it seems better to use instead of (1.5) the LOD method based on backward Euler, which has stronger damping properties and is also of 1-th order. Rather general convergence results for this LOD method were presented in [10].

For problems where more accuracy is demanded, the parallel alternating scheme (1.7) of [13], based on either (1.4) or (6.1), seems a good candidate. An alternative would be, for example, the Peaceman-Rachford ADI method. An analysis for this method can be found in [6]. Some numerical results given in [13] suggest that this ADI method and (1.7) are competitive.

A popular technique to improve accuracy is Richardson extrapolation. However, due to the fact that when both τ and h tend to 0, the structure of the temporal errors is different than for the classical ODE case $\tau \downarrow 0$, $h = h_0 > 0$. Therefore it is not clear yet whether extrapolation will increase the order of convergence in general. Some interesting results in this direction have been derived in [1] for the LOD method based on backward Euler. Numerical results can also be found in [15].

Acknowledgement. The author is grateful to J. Blom for providing the numbers of the Tables 5.1-5.3.

References

[1] W. Auzinger, On error structures and extrapolation for stiff systems with applications in the method of lines. Report Nr. 81/89, Inst. Appl. Numer. Math., Techn. Univ. Vienna, 1989.

[2] P. Brenner, M. Crouzeix and V. Thomée, Single step methods for inhomogeneous linear differential equations in Banach space. RAIRO Numer. Anal. 16 (1982), pp. 5-26.

[3] M. Fiedler, Special matrices and their applications in numerical mathematics. Martinus Nijhoff Publ., Dordrecht, 1986.

[4] P.J. van der Houwen and J.G. Verwer, One-step splitting methods for semi-discrete parabolic equations. Computing 22 (1979), pp. 291-309.

[5] W.H. Hundsdorfer and B.I. Steininger, Convergence of linear multistep and one-leg methods for stiff nonlinear initial value problems. Report NM-R8924, Centre Math. Comp. Sc., Amsterdam, 1989.

[6] W.H. Hundsdorfer and J.G. Verwer, Stability and convergence of the Peaceman-Rachford ADI method for initial-boundary value problems. Math. Comp. 53 (1989), pp. 81-101.

[7] P. Lancaster and M. Tismenetsky, The theory of matrices, second edition. Academic Press, Orlando, 1985.

[8] R.J. LeVeque, Intermediate boundary conditions for LOD, ADI and approximate-factorization methods. ICASE report 85-21, NASA Langley Research Centre, Hampton, Virginia, 1985.

[9] G.I. MARCHUK, Methods of numerical mathematics, second edition. Springer Verlag, New York, 1981.

[10] A.A. SAMARSKII, On an economical difference method for the solution of a multidimensional parabolic equation in an arbitrary region. Zh. Vych. Mat. 2 (1962), pp. 894-926.

[11] J.M. SANZ-SERNA, J.G. VERWER and W.H. HUNDSDORFER, Convergence and order reduction of Runge-Kutta schemes applied to evolutionary partial differential equations. Numer. Math. 44 (1987), pp. 405-418.

[12] B.P. SOMMEIJER, P.J. VAN DER HOUWEN and J.G. VERWER, On the treatment of time-dependent boundary conditions in splitting methods for parabolic differential equations. Int. J. Numer. Math. Engin. 17 (1981), pp. 335-346.

[13] D.A. SWAYNE, Time dependent boundary and interior forcing in locally one-dimensional schemes. SIAM J. Sci. Comp. 8 (1987). pp. 755-767.

[14] J.G. VERWER, Contractivity of locally one-dimensional splitting methods. Numer. Math. 44 (1984), pp. 247-259.

[15] J.G. VERWER and H.B. DE VRIES, Global extrapolation of a first order splitting method. SIAM J. Sci. Stat. Comp. 6 (1985), pp. 771-780.

[16] N.N. YANENKO, The method of fractional steps. Springer Verlag, Berlin, 1971.